

JANUSZ S. BIEN, ZYGMUNT SALONI

POJĘCIE WYRAZU MORFOLOGICZNEGO I JEGO ZASTOSOWANIE  
DO OPISU FLEKSJI POLSKIEJ

(wersja wstępna)

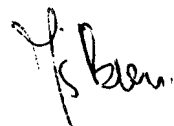
Celem niniejszego artykułu jest wyjaśnienie podstawowych znaczeń wieloznacznego terminu *wyraz*<sup>1</sup> oraz sformułowanie propozycji ścisłego opisu syntagmatycznych wyrazów morfologicznych (w pewnej mierze także wyrazów paradygmatycznych) w języku polskim. Jak wiadomo, ścisłość, zawsze wysoce pożądana, jest konieczna przy przygotowaniu materiału, w szczególności językowego, do badań maszynowych. Ujęcie problemu jest czysto językoznawcze, jednak ambicją autorów jest to, aby mogło ono być również punktem wyjścia do prac z zakresu lingwistyki informatycznej.

Artykuł w istotny sposób rozwija tezy sformułowane w innych pracach autorów<sup>2</sup> (zmiany w stosunku do tamtych prac wprowadzone są głównie ze względu na rozwinięcie aparatu pojęciowego). Nie ma on jednak charakteru ostatecznego, jako że poświęcony jest problemom będącym obecnie w fazie opracowania. W szczególności charakter tymczasowy ma stosowana w artykule terminologia. Autorzy nie wykluczają tego, że zmodyfikują ją w dalszych fazach pracy.

Tekst zasadniczo wymaga od czytelnika tylko ogólnej orientacji materiałowej w zakresie fleksji współczesnego języka polskiego oraz tradycyjnej terminologii fleksyjnej. Do jego pełniejszego zrozumienia może się przyczynić znajomość istniejących opisów fleksji polskiej. Czytelnik, który ją ma, bez trudu dostrzeże

<sup>1</sup> Wieloznaczność tego terminu (i jego odpowiedników w innych językach) była wielokrotnie analizowana w lingwistyce strukturalnej, por. np. J. Lyons, *Wstęp do językoznawstwa*, wyd. II, Warszawa 1973, s. 213—231; A. A. Зализняк, *Русское именное словоизменение*, Москва 1967, s. 19—22; J. Tokarski, *Fleksja polska*, wyd. II, Warszawa 1978, s. 15—24.

<sup>2</sup> Por. przede wszystkim: J. S. Bien, *O niektórych problemach przetwarzania języków fleksyjnych na maszynach cyfrowych*, *Prace Filologiczne* XXIII (1972), s. 187—191; J. S. Bien, *O dwóch pojęciach pożytecznych przy automatycznym przetwarzaniu tekstów*, (w:) *Z polskich studiów slawistycznych*, Seria 4, tom 1, *Językoznawstwo*, Warszawa 1972, s. 310—315; J. S. Bien, W. Łukaszewicz, S. Szpakowicz, *Opis systemu MARYSIA*, I. Zasady pisania scenariusza i scenopisu, *Sprawozdania Instytutu Maszyn Matematycznych i Zakładu Obliczeń Numerycznych Uniwersytetu Warszawskiego* nr 41, Warszawa 1973, s. 39; Z. Saloni, *Klasyfikacja gramatyczna leksemów polskich*, *Język Polski* LIV (1974), s. 3—13 i 93—101; Z. Saloni, *Kategoria rodzaju we współczesnym języku polskim*, (w:) *Kategorie gramatyczne grup imiennych w języku polskim*, Wrocław 1976, s. 43—116; Z. Saloni, *Kategorie gramatyczne liczebników we współczesnym języku polskim*, (w:) *Studia gramatyczne I*, Wrocław 1977, s. 145—173; Z. Saloni, *Uwagi o opisie fleksyjnym tw. zaimków rzeczownych* (w:) *Acta Universitatis Lodzensis, Folia Linguistica* 2, 1980, s. 243—253



inspirujący charakter prac Jana Tokarskiego<sup>3</sup> dla powstania niniejszego tekstu. Autorzy czasem przejmują z nich zaproponowane rozwiązania, często panujące już dziś w gramatyce polskiej, czasem są one punktem wyjścia ich własnych rozstrzygnięć. Rezygnują oni jednak z bezpośrednich odwołań do wymienionych prac, jak również innych prac, z których korzystają, ponieważ odwołania takie zajęłyby zbyt wiele miejsca.

Tekst pisany w języku naturalnym ma, ściśle rzecz biorąc, charakter dwuwymiarowy. Litery, pisane czy drukowane, są same przez się obiektami dwuwymiarowymi. Dwuwymiarowość odgrywa też istotną rolę przy wprowadzaniu takich rozróżnień specjalnych, jak wielkość i krój liter (czcionek). Kwestii tych możemy nie analizować.

Tekst pisany (w szczególności drukowany) w języku naturalnym należącym do naszego kręgu kulturowego zgodnie z ogólnie przyjętymi współczesnymi konwencjami składa się z wierszy, stanowiących ciągi wielkich i małych liter, znaków interpunkcyjnych i, ewentualnie, innych symboli. W sposób oczywisty, choć bynajmniej niełatwy do ścisłego opisania, w tekście drukowanym można wyróżnić jednostki, które konwencjonalnie nazwiemy **wyrazami typograficznymi**: są to ciągi liter (zarówno małych, jak i wielkich), ograniczone z obu stron znakiem interpunkcyjnym, odstępem lub końcem wiersza. Jednostki takie możemy, rzecz jasna, rozpatrywać jako okazy (tokens) i jako typy (types).

Zgodnie z powyższym określeniem, wyraz (rozumiany tradycyjnie) podzielony przy przeniesieniu jego części do nowej linii stanowi dwa wyrazy typograficzne, przy czym pierwszy z nich zawiera znak nie występujący w wyrazie (rozumianym tradycyjnie) dzielnym, a mianowicie łącznik<sup>4</sup>. Chociaż przeniesienie części wyrazów odbywa się zgodnie ze skodyfikowanymi regułami, podział tekstu na wyrazy typograficzne ma z punktu widzenia lingwistyki charakter przygodny. Lingwistycznego charakteru nie ma również większość wyróżnień drukarskich, do wyjątków należy może stosowanie kursywy w funkcji cudzośłowu. Z tego też powodu w dalszych rozważaniach abstrahujemy od typograficznych własności tekstu (w tym również od opozycji wielkich i małych liter) i traktujemy tekst — zgodnie zresztą z tradycją — jako twór jednowymiarowy. Jednostki powstałe w wyniku segmentacji tekstu jednowymiarowego (nie zawierającego przeniesień części wyrazu do nowej linii), dokonywanej za pomocą spacji, znaków interpunkcyjnych i łączników, nazywamy **wyrazami grafemicznymi** (mówiąc ściślej, są to odpowiednie ciągi elementów nazywanych grafemami, a reprezentujących litery alfabetu).

<sup>3</sup> Por. przede wszystkim: *Czasowniki polskie*, Warszawa 1951; *Z pogranicza metodyki i językoznawstwa*, Warszawa 1967; *Elementy gramatyki polskiej*, (w:) J. Stanisławski, *Wielki słownik polsko-angielski*, Warszawa 1969, s. 1559—1583; *Dialog: człowiek — maszyna cyfrowa, poszukiwanie wspólnego języka*, Prace Filologiczne XXIII (1972), s. 183—185; *Fleksja polska*, op. cit.

<sup>4</sup> W niniejszym tekście łącznik traktujemy jako literę. Nie uwzględniamy natomiast problemu nazwisk i nazw obcych, w których, zwłaszcza w przypadkach zależnych, występuje apostrof. Pomijamy również w zasadzie kwestię znaków przestankowych, wymagającą odrębnej analizy.

Pojęcie to jest bardzo przydatne i w praktyce, i w teorii. Z jednej bowiem strony wyrazy grafemiczne mogą być np. rozpoznawane automatycznie w tekstach zakodowanych na taśmie papierowej lub magnetycznej sterującej współczesnymi urządzeniami poligraficznymi, z drugiej zaś — nie zawierają elementów nieistotnych z lingwistycznego punktu widzenia. Wyrazy grafemiczne typy, traktowane jako zbiory wyrazów grafemicznych-okazów, są w zasadzie (z dokładnością do wspomnianych wyżej własności typograficznych) sumami teoriomnogościowymi wyrazów typograficznych-typów. Podstawę dalszych rozważań stanowić będą wyrazy grafemiczne-typy<sup>5</sup>.

Wyrazy grafemiczne języka polskiego nie są jednak jeszcze obiektem zainteresowania językoznawców ani nie są bezpośrednio notowane w słownikach. Dzieje się tak dlatego, że są one czasem jednostkami utworzonymi przez mechaniczne połączenie — na podstawie reguł ortograficznych — kilku jednostek mających bardziej podstawowy charakter. Wyrazy grafemiczne, które są połączeniami takich jednostek, zwłaszcza jednostek opisywanych w przepisach ortograficznych, gramatykach i słownikach jako formy wyrazów odmiennych albo wyrazy nieodmienne, z częstkami *nie* +, *+m*, *+em*, *+ś*, *+eś*, *+śmy*, *+eśmy*, *+eście*, *+ście*, *+by*(+), *+że*, *+ż*(+), *+li* będziemy nazywać aglutynatami<sup>6</sup>.

Po przejściu na wyższy poziom abstrakcji wprowadzimy pojęcie **wyrazu morfologicznego**, który jest — jeśli odwołamy się do aparatu językoznawstwa strukturalnego — jednostką bilateralną w odróżnieniu od poprzednio omawianych jednostek unilateralnych. Jego podstawowym elementem (w tradycji językoznawczej, *signifiant*) jest kształt grafemiczny, stanowiący pewien ciąg liter: będziemy go nazywać **słowem** (jest to świadome uproszczenie na użytek

<sup>5</sup> Pojęcie zbliżone do wprowadzonego tu wyrazu grafemicznego jest praktycznie stosowane w tych, obejmujących duże korpusy, pracach językoznawczych, które dążą do posługiwania się jednostkami danymi obiektywnie, a unikają jednostek powstałych w wyniku interpretacji danych. Słownik frekwencyjny języka polskiego operuje więc jako jednostkami podstawowymi słowoformami, które są dla nas wyrazami grafemicznymi (nie zinterpretowanymi bądź zinterpretowanymi) — tylko wyodrębnionymi nie w oryginalnym tekście drukowanym, lecz w tekście wtórnym, przepisany według specjalnych zasad (z zastąpieniem zapisów liczbowych słowami języka polskiego).

<sup>6</sup> Znak + wskazuje na miejsce przyłączenia ortograficznego. Jako aglutynatów nie traktujemy innych połączeń stosunkowo bliskich wymienionym. Pominęliśmy więc jednostki opisywane czasem jako połączenia form wyrazów odmiennych lub wyrazów nieodmiennych z częstkami: *+ś* (o znaczeniu nieokreśloności), *+kolwiek*, *+naście*, *+nastu*, *+nastoma*, *+dziesiąt*, *+dziesięciu*, *+dziesięcioma*, *+set*. Zasób takich połączeń jest bowiem stosunkowo niewielki, a tworzenie nie jest procesem żywym. Z tych samych względów do aglutynatów nie zaliczyliśmy połączeń przymków z częstką *+n*. Z powodu znikomej liczby wyrazy grafemiczne typu *doń*, *odeń* możemy potraktować jako kontrakcje (jak francuskie *au*, *aux*), mimo że cząstka *+n*, w przeciwieństwie do wyliczonych wyżej, na podstawie semantycznej i dystrybucyjno-składniowej daje się włączyć do paradygmatu. Częstki *+é*, jako potencjalnego elementu aglutynatów nie uwzględniliśmy z powodu ogromnej rzadkości oraz silnego ograniczenia stylistycznego jej użycia. Pominęte zostały też zrosty, nawet o obu członach odmiennych, jak *Białystok* czy *Wielkanoc* (które mogą być potraktowane jako konkatenaacje odpowiednich słów), oraz wszelkie złożenia oprócz tych, które pisze się z łącznikiem.

Możliwe jest również potraktowanie jako aglutynatów wyrazów grafemicznych, w których występują najbardziej produktywne przedrostki odmienne, zwłaszcza obcego pochodzenia, jak *anty* +, *pro* +, *eks* + itp. Kwestię tę traktujemy jako otwartą.

dalszych wywodów). Łącznik wewnątrz wyrazu grafemicznego (np. *angielsko-polski*) zawsze stanowi granicę słów. Jeśli wyraz grafemiczny lub jego część ograniczona łącznikiem stanowi aglutynat istniejącego wyrazu grafemicznego z jedną lub kilkoma spośród cząstek wymienionych wyżej<sup>7</sup>, to traktujemy go jako ciąg słów, stanowiących kształty odpowiednich wyrazów morfologicznych (dotyczy to również 1. i 2. osoby czasu przeszłego i trybu przypuszczającego). Jeśli wyraz grafemiczny (nie zawierający łącznika) stanowi aglutynat ciągu liter różniącego się od istniejącego wyrazu grafemicznego tylko regularną alternacją<sup>8</sup>, to traktujemy go również jako ciąg słów, stanowiących kształty różnych wyrazów morfologicznych (np. *mogłem* jako *mogl* + *+em*, *cóż* jako *có* + *+ż*). W pozostałych wypadkach każdy polski wyraz grafemiczny nie zawierający łącznika jest traktowany jako słowo stanowiące kształt jakiegoś wyrazu morfologicznego.

Wyraz morfologiczny oprócz kształtu posiada określone własności semantyczne i gramatyczne (w tradycji językoznawczej, *signifié*). Ponieważ na obecnym etapie rozważań interesują nas tylko opozycje morfologiczne, możemy założyć, że każdy wyraz morfologiczny wyposażony jest w znaczenie leksykalne, które może być reprezentowane przez pewien symbol<sup>9</sup>. Wyrazy morfologiczne zawierające ten sam symbol znaczenia leksykalnego zaliczane są do jednego wyrazu paradygmatycznego (termin ten nieco dokładniej skomentujemy niżej). Pozostałe elementy wyrazu morfologicznego będą miały postać indykatorów morfologicznych<sup>10</sup>. Są one dalej nazywane w sposób ułatwiający ich intuicyjną interpretację. Liczba indykatorów morfologicznych użytych dla różnych wyrazów morfologicznych będzie różna i zależeć będzie od stopnia komplikacji paradygmatu, do którego wchodzi dany wyraz morfologiczny. W zależności

<sup>7</sup> Istnieją tu oczywiście ograniczenia, i to łatwo dające się określić przez dystrybucję obu członów. Na przykład *kolegom* nie może być potraktowane jako aglutynat *kolego* i *+m*, podobnie *podeście* jako *pode* i *+ście*, *strzeliście* jako *strzeli* i *+ście*, *niebo* jako *nie* + *bo*, choć stanowią one konkatenacje odpowiednich ciągów liter. Ich ścisłe opisanie musiałoby się jednak opierać na określaniu łączliwości poszczególnych ciągów liter i byłoby bardzo kłopotliwe. Możliwa jest również homonimia wyrazów grafemicznych, typu *kogoś* (niepodzielne oraz *kogo* + *ś*). Rozróżnienie aglutynatu i wyrazu grafemicznego niepodzielnego na słowa może być jednak dokonane tylko przez odwołanie się do głębszych poziomów analizy tekstu: składniowego i semantycznego. Warto także zasygnalizować wypadki, gdy mimo podobnej podzielności na morfy aglutynat różni się od wyrazu grafemicznego niepodzielnego na słowa brakiem alternacji, np. *stółem* (*widział*) i *stołem* (*interesuje się*).

<sup>8</sup> I w tym wypadku, choć nielato, można znaleźć słowa homonimiczne, np. *pomiotłem* (niepodzielne oraz *pomiotł* + *+em*).

<sup>9</sup> Symbol ten, nazywany przez nas **reprezentantem paradygmatycznym**, w miarę możliwości powinien się pokrywać z tzw. formą hasłową czy formą słownikową wyrazu w sensie słownikowym. W wypadku homonimiczności form hasłowych dwu różnych wyrazów w dowolnym słowniku rozróżnia się je za pomocą kolejnych liczb. Podobnego zabiegu możemy dokonywać i w wypadku symboli reprezentantów paradygmatycznych.

<sup>10</sup> Termin ten, podobnie jak cała terminologia użyta w niniejszym artykule, ma charakter roboczy. Istotne w niniejszym tekście rozróżnienie cechy jako opozycji, możliwości przyjmowania pewnej właściwości (np. kolor) oraz cechy jako konkretnej właściwości (np. czerwony) jest w ogóle niezbyt mocno zakorzenione w polszczyźnie, także w jej stylu naukowym. Dlatego proponujemy terminy nie stosowane na ogół w pracach polskich i nie budzące niepożądanych skojarzeń.

od tej komplikacji wyróżnimy przede wszystkim poszczególne typy wyrazów morfologicznych, a więc wyrazy morfologiczne rzeczownikowe, wyrazy morfologiczne zaimkowe, wyrazy morfologiczne przymiotnikowe, wyrazy morfologiczne liczebnikowe, wyrazy morfologiczne czasownikowe, wyrazy morfologiczne pseudoimiesłowowe, wyrazy morfologiczne nieodmienne, dalej nazywane odpowiednio wyrazami rzeczownikowymi, wyrazami zaimkowymi itd. Ponieważ wyrazy czasownikowe tworzą zbiór pod względem fleksyjnym wyraźnie niejednorodny, zostanie on podzielony na podtypy fleksyjne za pomocą specjalnego indykatora *typ derywacyjno-fleksyjny*. Pozostałe typy wyrazów morfologicznych stanowią równocześnie typy fleksyjne (ale por. przypis 17).

We wszystkich wyrazach morfologicznych określonego typu lub podtypu fleksyjnego będą użyte te same obligatoryjne indykatory morfologiczne. W podzbiorach wyrazów morfologicznych określonego typu mogą ponadto, w miarę potrzeby, wynikającej z uwikłania poszczególnych słów w opozycje gramatyczne, wystąpić dodatkowo fakultatywne indykatory morfologiczne.

Dla poszczególnych typów wyrazów morfologicznych przewiduje się wyliczone niżej indykatory morfologiczne. Przy nazwach indykatorów fakultatywnych podajemy w nawiasie komentarz dotyczący zakresu ich użycia. Komentarz taki pomijamy wyłącznie w wypadku indykatora *variant*, mającego zawsze charakter fakultatywny. Wprowadzany on jest we wszystkich tych wypadkach, gdy w systemie języka znajdujemy różne słowa, realizujące — intuicyjnie rzecz biorąc — jednakowe funkcje semantyczne i syntaktyczne i skutkiem tego opisywane przez ten sam symbol znaczenia leksykalnego i te same wartości tych indykatorów morfologicznych, które odpowiadają kategoriom fleksyjnym. Indykator *variant* może być w miarę potrzeby dopisany do wszystkich wyrazów morfologicznych, także tych, w których nie został w niniejszej propozycji bezpośrednio uwzględniony. Po nazwie każdego indykatora po dwukropku podane są wartości, które może on przyjąć.

#### I. Wyrazy rzeczownikowe.

1. Liczba: *sing, pl.*<sup>11</sup>

2. Przypadek: *nom, gen, dat, acc, inst, loc, voc.*<sup>12</sup>

<sup>11</sup> Liczba traktowana jest jako cecha morfologiczna także w wypadku rzeczowników *plurale tantum* i *singulare tantum*. W wyrazie morfologicznym w żaden sposób nie jest sygnalizowana jego przynależność do paradygmatu defektywnego, w szczególności — paradygmatu nie obejmującego opozycji liczby, ponieważ nie należy to do charakterystyki słów jako jednostek. Informacje tego typu będą w słowniku podane automatycznie poprzez zasób wyrazów morfologicznych należących do wyrazu paradygmatycznego. Wobec braku opozycji fleksyjnej liczby charakteryzowanie słów, stanowiących kształty wyrazów morfologicznych należących do paradygmatów defektywnych, może odbywać się albo na podstawie kryterium składniowego, albo na zasadzie analogii kształtów graficznych rozważanego wyrazu paradygmatycznego z kształtami wyrazów paradygmatów niedefektywnych; w pierwszym wypadku rzeczowniki typu *wujostwo, generalostwo* kwalifikujemy jako *pl*, w drugim — jako *sing*. Sprawę tę pozostawiamy na razie otwartą.

<sup>12</sup> Uznanie wołacza za przypadek można uważać za konwencjonalne rozstrzygnięcie tradycyjne. Ze względu na głębsze piętra strukturalizacji zjawisk tekstowych, zwłaszcza składniowe, być może, dogodniejsze okaże się rozróżnienie wołacza i niewołacza za pomocą specjalnego indykatora.

3. Deprecjatywność (w wypadku *pl nom* rzeczowników, tradycyjnie zwanych męskoosobowymi): *neut* (*profesorzy* i *profesorowie*), *char* (*profesory*).

4. Apersonalność (w przypadku *pl nom* oraz *pl acc* pewnych rzeczowników): *neut* (*reprezentanci*, *reprezentantów*), *char* (*reprezentanty*)<sup>13</sup>.

5. Uniformizm (w wypadku *pl gen* niektórych rzeczowników): *neut* (np. *funkeji*), *char* (np. *funkeyj*).

6. Wariant: *var1* (np. *profesorowie*), *var2* (np. *profesorzy*)<sup>14</sup>. Należy tu uwzględnić między innymi następujące główne typy oboczności w deklinacji rzeczownika:

<i>sing gen</i>	: +a — +u	(karabina — karabinu),
<i>sing dat</i>	: +u — +owi	(kotu — kotowi),
<i>sing loc</i>	: +e — +u	(borze — boru),
<i>sing voc</i>	: +e — +u	(człowiecze — człowieku),
<i>pl nom</i>	: +i — +owie	(agronomi — agronomowie),
	+y — +owie	(profesorzy — profesorowie),
	+i — +e	(postaci — postacie),
	+y — +a	(gusty — gusta),
<i>pl gen</i>	: +i — +ów	(złodziei — złodziejów),
	+y — +ów	(pisarzy — pisarzy),
	+i — ∅	(sani — sań, alei — alej),
<i>pl inst</i>	: +ami — +mi	(gałęziami — gałęźmi).

## II. Wyrazy zaimkowe (w terminologii tradycyjnej: zaimki rzeczowne).

1. Liczba: *sing*, *pl*.<sup>15</sup>

2. Przypadek: *nom*, *gen*, *dat*, *acc*, *inst*, *loc*, *voc*.

3. Klityczność (w wypadku istnienia opozycji słów enklitycznych i nieenklitycznych): *neut* (słowo akcentowane, np. *mnie*, *jego*, *niego*), *char* (słowo nieakcentowane, np. *mi*, *mię*, *go*).

4. Poprzyimkowość (w niektórych wyrazach morfologicznych w przypadkach zależnych w formie akcentowanej): *neut* (forma niepoprzyimkowa, np. *jego*, *jemu*), *char* (forma poprzyimkowa, np. *niego*, *niemu*).

5. Aglutynacyjność (w wypadku zróżnicowania słów występujących samodzielnie i w aglutynatach): *neut* (np. *kto*, *co*, *kogo*, *czego*), *char* (np. *któ+*, *có+*, *kogó+*, *czegó+*)<sup>16</sup>.

<sup>13</sup> Warto zwrócić uwagę na to, że różnym wartościom indykatora *deprecjatywność* i *apersonalność* odpowiadają różne wartości rodzaju selektywnego rzeczownika; tak więc kategorie te mają charakter morfologiczno-składniowy, a w niektórych przypadkach nawet czysto składniowy (np. *ci nauczyciele byli* — *te nauczyciele były*).

<sup>14</sup> Decyzja, który z wariantów uznamy za pierwszy (zasadniczy), który — za drugi, jest w istocie arbitralna, choć powinniśmy dążyć do tego, by wiązała się ona ze wskazówką normatywną. Wymaga to jednak sprecyzowania kryteriów oceny, np. w postaci przewagi statystycznej formy zakwalifikowanej jako zasadnicza nad drugą w korpusie tekstów. Aby otrzymać takie wyniki, trzeba dysponować ogromnymi korpusami tekstów dostępnymi dla maszyny. To jednak wydaje się sprawą stosunkowo bliskiej przyszłości.

<sup>15</sup> Indykator *liczba* będzie w tym wypadku określał cechy nie morfologiczne, lecz składniowe.

<sup>16</sup> Osobnym zagadnieniem do opracowania jest dystrybucja formy aglutynacyjnej i nieaglutynacyjnej (por. np. *Coś zrobił?*, ale *Coż zrobił?*).

### III. Wyrazy przymiotnikowe<sup>17</sup>.

1. Stopień (dla przymiotników stopniujących się): *pos, comp, sup*.
2. Liczba: *sing, pl*.
3. Przypadek: *nom, gen, dat, acc, inst, loc, voc*.
4. Rodzaj: *m1* (męskoosobowy), *m2* (męski żywotny nieosobowy), *m3* (męski nieżywotny), *f, n1, n2* (zróżnicowanie rodzaju nijakiego istotne tylko dla konstrukcji z liczebnikami — por. niżej)<sup>18</sup>.
5. Kontraktywność (dla tzw. tradycyjnie zaimków dzierżawczych): *neut* (np. *moim, swoje*), *char* (np. *mym, swe*).
6. Predykatywność (dla nielicznych wyrazów o wartości *posing nom m*): *neut* (np. *gotowy*), *char* (np. *gotów*).
7. Wariant: *var1* (np. *pospieszny*), *var2* (np. *pośpieszny*).

### IV. Wyrazy liczebnikowe.

1. Przypadek: *nom, gen, dat, acc, inst, loc, voc*.
2. Rodzaj: *m1, m2, m3, f, n1* (przypisywany słowom nazywanym tradycyjnie liczebnikami zbiorowymi, np. *dwoje, dwojga, dwojgu, dwojgiem*), *n2* (przypisywany pozostałym słowom liczebnikowym łączącym się z tzw. rzeczownikami nijakimi,<sup>19</sup>
3. Akomodacyjność (istotna w wypadku *nom m1* pięciu par słów): *congr* (wyraz morfologiczny narzucający związanemu z nim słowu rzeczownikowemu wartość przypadku równą własnej, np. *dwaj*), *rec* (wyraz morfologiczny narzucający związanemu z nim słowu rzeczownikowemu wartość przypadku różną od własnej, np. *dwóch, dwu*).
4. Wariant (w wypadku dubletów dla ustalonych wartości innych indyktorów): *var1* (np. *dwóch, pięciu*), *var2* (np. *dwu, pięcioma*).

### V. Wyrazy czasownikowe.

System koniugacyjny języka polskiego jest, jak wiadomo, skomplikowany. W skład paradygmatu czasownikowego wchodzi słowa pozostające niewątpliwie w regularnych stosunkach morfologicznych, lecz nie przeciwstawiane sobie wzajemnie na jednolitych zasadach. Obok słów, którym przypisuje się określoną wartość morfologiczną osoby, znajdują się w nim słowa, którym nie można przypisać żadnej wartości osobowej. Są to w szczególności słowa układające się wewnątrz paradygmatu czasownikowego w pełny paradygmat

<sup>17</sup> Wydaje się pożądane uwzględnienie w paradygmacie przymiotnika formy zwanej tradycyjnie przysłówkiem, a w konsekwencji wprowadzenie indykatora *typ derywacyjno-fleksyjny* (patrz punkt V) dla wyrazu przymiotnikowego. Konkretnie wartości dla tego indykatora proponujemy jednak dopiero po rozważeniu kwestii, czy nie uwzględnić — *mutatis mutandis* — również rzeczowników odprzymiotnikowych na +ość.

<sup>18</sup> Ponieważ w paradygmatach wyrazów różnych typów fleksyjnych występuje systematyczna homonimiczność względem pewnych rozróżnień rodzajowych, w praktyce wygodniej będzie posługiwać się wartościami tego indykatora w postaci tak dobranych symboli złożonych, aby odpowiednie elementy ich struktury wskazywały bezpośrednio na tradycyjne kategorie osobowości, żywotności itp.

<sup>19</sup> Słowa stanowiące kształty wyrazów liczebnikowych odznaczają się na ogół znaczną homonimicznością. Na przykład słowa *pięć* i *pięciu* odpowiadają wszystkim wyrazom morfologicznym o wartości rodzajowej *m1, m2, m3, f i n2* oraz, ewentualnie, oznaczonych wartością *var1* (w wypadku wartości *inst* indykatora *przypadek*).

(lub podparadygmat) liczbowo-przypadkowy (odśłownik) lub liczbowo-przypadkowo-rodzajowy (imiesłowy przymiotnikowe), a także słowa, których specjalna budowa i funkcja przeciwstawia je wszystkim pozostałym słowom tego samego paradygmatu (bezosobnik, bezokolicznik, imiesłów przysłówkowy). Dlatego w istocie paradygmat czasownikowy zostaje rozbity na oparte na różnych kombinacjach opozycji podparadygmaty: przynależność wyrazu morfologicznego do podparadygmatu sygnalizowana jest za pomocą specjalnego indykatora *typ derywacyjno-fleksyjny*. Wszystkie wyrazy morfologiczne czasownikowe mają dwa indykatory wspólne: *aspekt* (którego rola w opisie zostanie scharakteryzowana niżej) oraz *typ derywacyjno-fleksyjny*. Zasób pozostałych indykatorów obligatoryjnych zależy od wartości indykatora *typ derywacyjno-fleksyjny*, wyznaczającej podtyp fleksyjny. Niżej zostaną jednak najpierw wyliczone wszystkie indykatory wraz z wartościami, potem zaś scharakteryzowane ich wykorzystanie dla poszczególnych podtypów fleksyjnych.

1. Aspekt: *imperf*, *perf*.

2. Typ derywacyjno-fleksyjny:

a. predyktor zwykły (obejmujący słowa, będące wykładnikami czasu teraźniejszego lub przyszłego prostego),

b. predyktor przyszły (tylko dla słów *będę*, *będziesz*, *będzie*, *będziemy*, *będziecie*, *będą*),

c. predyktor aglutynacyjny (tylko dla słów  $+m$ ,  $+em$ ,  $+ś$ ,  $+eś$ ,  $+śmy$ ,  $+eśmy$ ,  $+ście$ ,  $+eście$ )<sup>20</sup>,

d. rozkaznik,

e. pseudoimiesłów (słowa typu *pisał*, ewentualnie uwikłane w połączenia typu *będę pisał*, *pisałem*, *kiedym pisał*),

f. bezosobnik (słowa typu *pisano*),

g. bezokolicznik,

h. imiesłów przysłówkowy (w wypadku aspektu *imperf* na  $+qc$ , w wypadku aspektu *perf* —  $+szy$ ),

i. przymiotnik odczasownikowy bierny (imiesłów przymiotnikowy bierny),

j. przymiotnik odczasownikowy czynny (w wypadku wartości aspektowej *imperf* — na  $+qc$ , w wypadku wartości aspektowej *perf* — na  $+l$ , jeżeli istnieje)<sup>21</sup>,

k. rzeczownik odczasownikowy (zwany też odśłownikiem i gerundium).

3. Liczba: *sing*, *pl*.

4. Osoba: *prim*, *sec*, *ter*.

5. Przypadek: *nom*, *gen*, *dat*, *acc*, *inst*, *loc*, *voc*.

6. Rodzaj: *m1*, *m2*, *m3*, *f*, *n1*, *n2*.

<sup>20</sup> Słowa *jestem*, *jestes*, *jestesmy*, *jestescie* traktujemy jako niepodzielne. Otrzymują one wartość typu derywacyjno-fleksyjnego: predyktor zwykły.

<sup>21</sup> W sposób analogiczny jak w przypadku przymiotników właściwych do podparadygmatu przymiotnika odczasownikowego czynnego można włączyć słowa typu *śpiewająco*, *wyzywająco*.



7. Wokaliczność: *neut* (+*m*, +*ś*, +*śmy*, +*ście*), *char* (+*em*, +*eś*, +*eśmy*, +*ęcie*)<sup>22</sup>.

8. Aglutynacyjność: *neut* (np. *mókt*, *secht*), *char* (np. *mokt* +, *secht* +)<sup>23</sup>.

9. Wariant: *var1*, *var2*, *var3* (np. *szezcze*, *szezcze*, *szepta*).

Indykator *osoba* (wraz z indykátorem *liczba*) używany jest tylko do charakterystyki wyrazów morfologicznych o typie derywacyjno-fleksyjnym *a*, *b*, *c* i *d*. W wypadku predykatora zwykłego i przyszłego występuje pełny zasób wartości osobowych; w wypadku predykatora aglutynacyjnego oraz rozkaźnika o wartości liczbowej *pl* — wartości *prim* i *sec*; w wypadku rozkaźnika o wartości liczbowej *sing* — tylko wartość *sec*. Ponadto dla predykatora aglutynacyjnego stosowany jest indykator obligatoryjny *wokaliczność*, odróżniający słowa występujące po samogłosce i po spółgłosce.

Do dalszej charakterystyki pseudomiesłowu używa się indykatorów obligatoryjnych: *liczba* i *rodzaj*. Indykator fakultatywny *aglutynacyjność* wprowadzany jest w miarę potrzeby w tych wypadkach, gdy ma miejsce zróżnicowanie słów w zależności od kontekstu w wyrazie graficznym (por. przykłady wyżej).

Podtypy fleksyjne przymiotników odczasownikowych biernych i przymiotników odczasownikowych czynnych opisywane są podobnie jak typ fleksyjny przymiotników. Ich charakterystyka zatem zawiera indykatory obligatoryjne: *liczba*, *przypadek*, *rodzaj*, podobnie dalsza charakterystyka rzeczowników odczasownikowych — indykatory obligatoryjne: *liczba*<sup>24</sup> i *przypadek*. Charakterystyka morfologiczna bezosobnika, bezokolicznika i imiesłowu przysłówkowego nie zawiera żadnych dodatkowych indykatorów obligatoryjnych i ogranicza się do aspektu, typu derywacyjno-fleksyjnego i — ewentualnie — wariantu.

Indykator *wariant* jest użyty w tych wszystkich wypadkach, gdy istnieją warianty swobodne przy ustalonych wartościach pozostałych indykatorów, np.:

*przekonuje* — *przekonywa*<sup>25</sup>, *cieknie* — *ciecze*, *szezcze* — *szezcze* — *szepta*,

*mókt* — *moknął*,

*weź* — *weźmij*,

*biec* — *biegnąć*, *kaszać* — *kaszeć* — *kasać*,

*przekonując* — *przekonywując*<sup>25</sup>,

*przekonujący* — *przekonywający*<sup>25</sup>,

*przekonującego* — *przekonywającego*<sup>25</sup>.

<sup>22</sup> Dystrybucja wyrazów morfologicznych obu grup jest uwarunkowana fonologicznie i zupełnie jasna. Wyrazy morfologiczne o wartości *neut* występują po samogłosce, o wartości *char* — po spółgłosce.

<sup>23</sup> Warto zwrócić uwagę na to, że dystrybucja przeciwstawionych tu form zależy rzeczywiście od kontekstu w wyrazie graficznym i nie może być, jak często podają gramatyki, wiązana z określonymi wartościami osobowo-liczbowo-rodzajowymi (np. *abym mókt*, *abym secht*).

<sup>24</sup> Regularnym rzeczownikom odczasownikowym możemy nie tylko przyporządkowywać wartość liczbową na zasadzie składniowej, lecz nawet postulować istnienie tu opozycji fleksyjnej.

<sup>25</sup> Warto zwrócić uwagę na to, że w istocie pary słów opatrzone odnośnikiem 25 dają wariant znacznej części paradygmatu. Budowa morfemowa serii zróżnicowanych słów sugerowałaby tu nawet rozróżnienie symboli znaczenia leksykalnego PRZEKONYWAĆ 1 (dla *przekonuje*, *przekonujący*) i PRZEKONYWAĆ 2 (dla *przekonywa*, *przekonywający*), a także ewen-

VI. Wyrazy pseudoimiesłowowe samodzielne (*powinien, winien, rad* <sup>26</sup>).

1. Liczba: *sing, pl*.

2. Rodzaj: *m1, m2, m3, f, n1, n2*.

VII. Wyrazy nieodmienne, których opis nie zawiera obligatoryjnych indykatorów morfologicznych.

1. Wokaliczność (dla niektórych przyimków): *neut* (np. *w, przed*), *char* (np. *we, przede*) <sup>27</sup>.

2. Aglutynacyjność (dla bardzo nielicznych wyrazów): *neut* (np. *dla czego*), *char* (np. *dla czego* w wyrazie graficznym *dla czegoż*).

3. Wariant: *var1* (np. *piure*), *var2* (np. *purée*).

Do grupy tej zostają, na czysto formalnej zasadzie, zaliczone wszystkie słowa nie przeciwstawione innym za pomocą kategorii gramatycznych uwzględnionych w opisie wyliczonych wyżej klas. Potraktowane są w ten sposób słowa różnofunkcyjne i mające rozmaite właściwości składniowe. W szczególności warto zwrócić uwagę na przynależność do tej grupy następujących klas słów:

a. Słowa odpowiadające wyróżnianym tradycyjnie tzw. częściom mowy nieodmiennym oprócz regularnych przysłówków odprzymiotnikowych, które zostaną zaliczone do paradygmatów przymiotnikowych.

b. Słowa interpretowane jako części mowy odmienne, lecz mające pełną neutralizację wszystkich form, tj. rzeczowniki nieodmienne, np. *foyer* (w tym skróty i skrótownice, np. *zsrr*), przymiotniki nieodmienne, np. *khaki*, czasowniki nieodmienne syntetycznie, np. *można, trzeba*.

Należy zauważyć, że pojęcie wyrazu morfologicznego nie pokrywa się z pojęciem **formy wyrazowej** w rozumieniu tradycyjnym. Ze względów technicznych za punkt wyjścia w tworzeniu wyrazów morfologicznych przyjmujemy wyrazy grafemiczne, spośród których przez podział wyeliminowaliśmy jednostki utworzone w sposób czysto przypadkowy (na podstawie reguł ortograficznych). Słowa stanowiące kształty wyrazów morfologicznych są więc zawsze jednolitymi (nie rozerwanymi) i nie zawierającymi spacji ciągami liter. Jako słowa nie są zatem traktowane tzw. formy analityczne, np. *będę pisał*, albo jednostki typu *boję się*, które bywają i mogą być uważane za wyrazy (czy „formy wyrazów”), lecz niewątpliwie w innym sensie. W sposób podobny do form analitycznych są przez nas traktowane, skutkiem przeprowadzonych podziałów, niektóre wyrazy grafemiczne tradycyjnie opisywane w morfologii jako formy syntetyczne. Kształtem graficznym odpowiedniego wyrazu morfologicznego jest natomiast każde ze słów składających się na formę analityczną <sup>28</sup>.

tualnie PRZEKONYWAĆ 3 (dla *przekonywuje, przekonywujący*); rozwiązanie takie zbyt odbiegłoby jednak od tradycji.

<sup>26</sup> O ile nam wiadomo, podane słowa reprezentują wszystkie paradygmaty zbudowane wyłącznie na opozycji liczby i rodzaju. Listę można ewentualnie uzupełnić archaicznym *koutent*.

<sup>27</sup> Dystrybucja członów dwóch spośród opozycyjnych par (*w — we, z — ze*) jest dosyć skomplikowana i wymaga dodatkowych badań, dystrybucja pozostałych ma uzasadnienie czysto leksykalne. Do wymienionej grupy można ewentualnie włączyć parę słów *+z — +ze* (+).

<sup>28</sup> Podobne zasady były przyjmowane także przez innych badaczy (por. H. Misz, *Opis grup syntaktycznych dzisiejszej polszczyzny pisanej*, Bydgoszcz 1967; И. А. Мельчук, *Опыт теории лингвистических моделей типа «смысл ↔ текст»*, Москва 1974).

Wyraz morfologiczny traktowany jest jako zbiór  $n$ -elementowy. Liczba jego elementów jest różna dla różnych wyrazów morfologicznych, jednak wynosi co najmniej 2 (dla niektórych wyrazów nieodmiennych, reprezentowanych tylko przez kształt graficzny i symbol znaczenia leksykalnego). Indykatory morfologiczne wprowadzone na dalszych miejscach powinny objąć wszystkie<sup>29</sup> rozróżnienia słów występujące wewnątrz wyrazów słownikowych, wyodrębnionych w opisie języka polskiego zgodnie z tradycją i intuicją, a ponadto inne najregularniejsze ich opozycje, w których można się dopatrywać żywych mechanizmów czy operacji. W wypadkach takich postuluje się przeciwstawienie obu członów opozycji w sposób stosowany dla faktów gramatycznych, bez zwiększania liczby elementów leksykalnych (rozróżnianych przy pomocy symbolów znaczenia leksykalnego). W niniejszej propozycji zostało to osiągnięte w dwojaki sposób. Po pierwsze, za fleksyjne uznano część opozycji, tradycyjnie traktowanych inaczej (w ten sposób przeciwstawiono np. tzw. liczebniki zbiorowe — tzw. liczebnikom głównym). Po drugie (w tym wypadku niemal całkowicie zgodnie z tradycją), ze względów praktycznych przeciwstawiono sobie wzajemnie słowa składniowo i morfologicznie dalekie, lecz dające się pogrupować w serie regularne, choć skomplikowane, jak jest w wypadku słów czasownikowych należących do różnych podtypów fleksyjnych.

Wprowadzenie indykatorów oraz ich wartości do całej klasy wyrazów morfologicznych opiera się w istocie na kompromisie między potencjalną liczbą opozycji, w których uczestniczą słowa, stanowiące kształty wyrazów morfologicznych należących do danej klasy, a faktycznym istnieniem wykładników opozycji. W każdym razie, jeśli za różnym traktowaniem tego samego słowa przemawiają — oprócz istnienia podobnej opozycji morfologicznej o co najmniej dwu różnych wykładnikach w innych klasach wyrazów — względy składniowe, to traktujemy je jako słowo homonimiczne, stanowiące kształt dwu różnych wyrazów morfologicznych. Tak więc, zgodnie z tradycją, przyjmujemy, że każdemu przypadkowi rzeczownika czy liczebnika (przy ustalonych wartościach pozostałych indykatorów) odpowiada inny wyraz morfologiczny. Decyzja taka ułatwi wprowadzenie do opisu jednostek leksykalnych właściwości składniowych, które będą w znacznym stopniu pokrywały się z wprowadzonymi wcześniej indykatorami morfologicznymi. Jeśli natomiast potencjalnej, lecz zneutralizowanej opozycji morfologicznej nie towarzyszy różnicowanie właściwości składniowych, możemy danego indykatora morfologicznego nie wprowadzać. Wpływają na to nie względy teoretyczne, lecz techniczne. Konsekwentniejsze byłoby niewątpliwie wprowadzenie do opisu — zgodnie z postulatami lingwistyki strukturalnej — wszystkich opozycji potencjalnie rozróżnialnych, spowodowałoby ono jednak znaczne zwiększenie liczby jednostek podlegających opisowi.

<sup>29</sup> Nie wykluczamy jednak tego, żeśmy mogli coś przeoczyć. Jeśli w trakcie opracowywania materiału zostaną znalezione w języku polskim opozycje morfologiczne nie uwzględnione w powyższym zestawie, trzeba będzie zmodyfikować zasób lub wartości indykatorów, za pomocą których opisuje się daną jednostkę.

Sposób rozszerzenia opisu o indykatory czy parametry składniowe wymaga osobnej dyskusji. Najpierw powinno być przedyskutowane samo pojęcie **wyrazu składniowego**. Szczególnie trzeba wziąć pod uwagę to, iż jeśli postępować zgodnie z polską tradycją gramatyczną i ukształtowanymi na niej intuicjami, za wyrazy składniowe (tj. jednolite, nieredukowalne składniki struktury zdania) wypada uznać ciągi (być może, poprzedzielane innymi elementami) wyrazów morfologicznych, np. *biało-czerwony*, *angielsko-rosyjski*, *po angielsku*<sup>30</sup>, *po cichu*, *będę pisać*, *będę pisał*, *+m pisał*, *+by+ +m pisał*, *boję się*, *w lesie*, *na schwał*, *z nienacka*<sup>31</sup> itd. Przy definiowaniu pojęcia wyrazu składniowego konieczne jest ściśle pogrupowanie takich wypadków w serie i podjęcie dla każdej z osobna decyzji. Warto zwrócić uwagę na to, że w wypadku wyrazów składniowych składających się z kilku wyrazów morfologicznych niewykłuczone jest wprowadzenie parametrów składniowych dla całości. Wartość konkretnego parametru składniowego wyrazu składniowego może być wtedy pochodna wobec indykatora morfologicznego którejś z jego części.

W każdym razie nawet w wypadkach, gdy wyraz morfologiczny i wyraz składniowy są reprezentowane przez ten sam kształt, liczba parametrów wyrazu składniowego powinna być — z reguły — większa niż wyrazu morfologicznego. Wynika to w sposób oczywisty już z przedstawionego wyżej opisu. Na przykład wyraz morfologiczny rzeczownikowy nie jest w żaden sposób scharakteryzowany co do rodzaju: cecha ta musi być uwzględniona w charakterystyce składniowej.

Zaproponowany wyżej opis wyrazów morfologicznych prowadzi oczywiście do pogrupowania słów w sposób poprawny lingwistycznie i, w miarę możliwości, zgodny z polską tradycją gramatyczną i słownikową<sup>32</sup>. Pogrupowanie to następuje w wyniku utworzenia zbiorów wyrazów morfologicznych o identycznym symbolu znaczenia leksykalnego (reprezentanta paradygmatycznego — por. przypis 9). Zbiory takie nazwiemy **wyrazami paradygmatycznymi**. Wyrazy morfologiczne należące do tego samego wyrazu paradygmatycznego różnią się między sobą kształtem (wyraz paradygmatyczny może być więc reprezentowany przez różne słowa) oraz wartościami indykatorów morfologicznych, a czasem i ich zasobem.

Zwróćmy uwagę na to, że nie wszystkie indykatory morfologiczne opisują cechy, które skłonni byśmy uważać za fleksyjne. W szczególności indykator *typ derywacyjno-fleksyjny* wprowadza rozróżnienia, naszym zdaniem, po części słowotwórcze, po części — fleksyjne (brak konsekwencji jest tu spowodowany technicznymi trudnościami opisu bardzo skomplikowanego materiału). Zatem

<sup>30</sup> Trzeba tu zwrócić uwagę na słowa *angielsko*, *angielsku*. Sposób uwzględnienia takich słów w opisie morfologicznym języka polskiego i włączenia do wyrazów paradygmatycznych typu ANGIELSKI zostanie przedyskutowany w toku dalszej pracy.

<sup>31</sup> Do dyskusji i rozstrzygnięcia pozostaje sposób opisu i, ewentualnie, włączenia do wyrazów paradygmatycznych wszelkich rzadkich słów reliktowych i frazeologicznie uwikłanych.

<sup>32</sup> Postulat zgodności z tradycją jest wcale niebłaży. Rozwiązania tradycyjne ułatwiają dostęp do informacji nie przeszkolonemu człowiekowi, wpływają na łatwość akceptacji opisu w praktyce badawczej i dydaktycznej.

wyraz paradygmaticzny czasownikowy — i ewentualnie przymiotnikowy — obejmuje różne jednostki, choć ściśle spokrewnione słotwórczo. Wyraz paradygmaticzny tego rodzaju proponujemy nazywać **wokabulą**, podczas gdy wyraz paradygmaticzny o jednorodnym pod względem fleksji charakterze — a także analogiczny podparadygmat wokabuly — można nazywać np. **fleksemem**. W niniejszym tekście staraliśmy się opisać przede wszystkim wyrazy typu fleksmu, nie przesadzając sprawy organizacji wokabul. W szczególności nie sprecyzowaliśmy tego, czy wyrazom czasownikowym różniącym się aspektem przypisywać zawsze różne symbole znaczenia leksykalnego, czy czasem taki sam (należałyby wówczas do jednej wokabuly). Nawet w wypadku pierwszej decyzji indyktor *aspekt* jest jednak potrzebny, ponieważ różnice między wyrazami paradygmaticznymi o aspekcie niedokonanym i dokonanym wymagają notowania odpowiedniej informacji w słowniku.

Przez **paradygmat** rozumiemy tutaj ustalony porządek elementów wyrazu paradygmaticznego<sup>33</sup>; analogicznie, ustalony podzbiór wyrazu paradygmaticznego, uporządkowany według określonych zasad, możemy nazywać **podparadygmatem**. W wypadku wprowadzenia zaproponowanych tu pojęć do systematycznego wykładu gramatyki polskiej konieczne będzie ściśle sformułowanie zasad porządkowania elementów poszczególnych typów wyrazów paradygmaticznych i wybierania podparadygmatów. Celem niniejszego tekstu nie jest jednak dyskusowanie zasad organizacji gramatyki, lecz wyłącznie zaproponowanie aparatu pojęciowego dla jej potrzeb.

Odrębno potraktowania wymaga pojęcie **leksemu**. W pracach językoznawczych przez leksemy rozumie się na ogół podstawowe elementy słownictwa danego języka, jednak szczegółowa interpretacja postaci tych elementów bywa bardzo różna. Ostatnio coraz częstsze, zarówno w pracach zagranicznych, jak i polskich, staje się rozumienie leksemu jako zbioru pewnych jednostek powstałych w wyniku interpretacji jednostek tekstowych. Podejście takie jest nam bliskie, jednak problem szczegółowej reprezentacji leksemów pragniemy pozościć poza dyskusją. Przy definiowaniu leksemów powinien być także wzięty pod uwagę czynnik semantyczny, dokładniej zanalizowany i określony niż znaczenie leksykalne sygnalizowane w wyrazach morfologicznych przez odpowiedni reprezentant paradygmaticzny<sup>34</sup>.

<sup>33</sup> Właściwie trzeba by tu mówić o paradygmatach morfologicznych. Możliwe jest bowiem tworzenie paradygmatów wyrazów składniowych (a także innych jednostek). W szczególności tablice odmiany podawane przez podręczniki gramatyki są właśnie paradygmatami wyrazów składniowych (przeciwstawiają sobie na przykład jednostki typów *piszę* i *będę pisać*).

<sup>34</sup> Łatwo podać przykłady ilustrujące, że proponowany tu aparat nie wystarcza do wyodrębniania leksemów. Przede wszystkim rozróżnienie dwóch identycznych serii słów, różniących się jaskrawo znaczeniem (np. *pokój*, *pokoju*, ...) musi bazować na intuicjach znacznie wykraczających poza zjawiska zaliczane do tradycyjnych działów gramatyki, morfologii i składni. Nawet dwie serie słów, z których jedna jest podzbiorem drugiej, rozróżnić na podstawie znaczenia jest łatwo tylko w szczególnych wypadkach, najczęściej wyrazów paradygmaticznych rzeczownikowych (np. *zwłoka*, ..., *zwłocę*, *zwłoki*, ..., *zwłokach* oraz *zwłoki*, ..., *zwłokach*). Zauważenie podobnych stosunków w innych wypadkach wymaga co najmniej subtelnej analizy składniowej. Przykładami mogą być leksemy, do których zostaną zaliczone słowa wyróżnione w zdaniach:

Istnieje oczywiście potrzeba operowania jednostkami słownictwa języka naturalnego roboczo, w sposób doraźny, przy tworzeniu kompendiów w rodzaju słowników jednojęzycznych i dwujęzycznych oraz posługiwaniu się nimi. Taką jednostką roboczą, jest **wyraz słownikowy**, reprezentowany w samym słowniku oraz w rozmaitych opracowaniach przez **słowo hasłowe**. Wyraz słownikowy stanowi często przybliżenie leksemu, jednak nie jest to regułą. Możliwe jest tworzenie haseł gniazdowych, obejmujących po kilka leksemów połączonych najregularniejszymi stosunkami derywacyjnymi, jak proponujemy w wypadku wyrazów paradygmatycznych czasownikowych i — ewentualnie — przymiotnikowych (wyraz słownikowy jest w takich wypadkach wokabulą). Określenie wyrazu słownikowego, podobnie jak wszystkich jednostek omawianych w niniejszym tekście, powinno zostać wyprowadzone z konkretnych obiektów fizycznych. Elementami tymi są w naszym wypadku znaki drukarskie, mające z natury charakter dyskretny, w przeciwieństwie do ciągłych fal głosowych, stanowiących tworzywo tekstu mówionego.

Wszystkie wyliczone wyżej obiekty: wyrazy typograficzne, wyrazy graficzne, wyrazy morfologiczne (i reprezentujące je słowa), wyrazy składniowe (wraz z odpowiednią reprezentacją), leksemy, fleksemy i wokabuly, wyrazy słownikowe (i reprezentujące je słowa hasłowe) w opisie języka są w istocie definiowane przez słowniki, tj. odpowiednie listy. Chociaż w dokładnym opisie języka listy takie powinny być podane w sposób jawny, nie zawsze wymaga to ręcznego sporządzania dużych spisów o takim charakterze, bo w wielu wypadkach wystarcza ścisła instrukcja przetwarzania jednej listy (np. spełniającego odpowiednie warunki spisu wyrazów słownikowych) w drugą<sup>35</sup>.

Zresztą wobec właściwości słownika, zwanej potocznie otwartością, listy takie mają zawsze charakter niepełny i mogą być kontynuowane. Kontynuacja powinna się odbywać na zasadzie regularności, która jest podstawą wyodrębniania wszystkich elementów uwzględnionych na listach.

Świadomość odrębności wszystkich pojęć wprowadzonych wyżej powinna znaleźć odbicie w opisie gramatycznym współczesnego języka polskiego<sup>36</sup>. Powinna ona stanowić również istotny element przy opracowaniu założeń metodologicznych nowego ambitnego słownika języka polskiego, stawiającego sobie cele praktycznej przydatności i ścisłości odpowiadającej dzisiejszemu poziomowi wiedzy językoznawczej<sup>37</sup>. Na wstępnym etapie pracy nad takim

*Jan należy do partii. — Należy uważać.*

*Jan zapisał się do partii. — Tym piórem pisze się łatwo.*

*Jestem gotów do wyjścia. — Jużem gotów wyjść.*

<sup>35</sup> Za taką ścisłą instrukcją przetwarzania listy słów hasłowych wyrazów słownikowych w listę słów stanowiących kształty wyrazów morfologicznych może być uważany słownik gramatyczny języka rosyjskiego A. A. Zalizniaka (Moskwa 1977).

<sup>36</sup> Podobnie konieczne jest zanalizowanie pojęć potrzebnych do opisu zjawisk brzmieniowych.

<sup>37</sup> Ogromną zasługą Jana Tokarskiego jest opracowanie koncepcji informacji gramatycznej w jedenastotomowym *Słowniku języka polskiego* PAN pod red. W. Doroszewskiego. Koncepcja

słownikiem konieczne byłoby przedyskutowanie zasad zaliczania elementów tekstowych do wyrazów słownikowych. Rozstrzygnięcia powinny być dokonane na podstawie jasnej koncepcji opisu gramatycznego języka polskiego <sup>38</sup>.

ta, stworzona przed ćwierć wiekiem, jest w pełni adekwatna do stanu wiedzy językoznawczej z owego czasu i wyraźnie góruje nad analogicznymi elementami słowników innych języków słowiańskich.

<sup>38</sup> Autorzy dziękują wszystkim, którzy przekazali im swoje uwagi o wcześniejszych wersjach niniejszego tekstu, szczególnie kol. kol. Włodzimierzowi Gruszczyńskiemu, Marii Szupryczyńskiej, Markowi Świdzińskiemu i Adamowi Weinsbergowi.